Structural Organization of the WD repeat protein-encoding gene HIRA in the DiGeorge syndrome critical region of human chromosome 22.

S Lorain, S Demczuk, V Lamour, et al.

Genome Res. 1996 6: 43-50
Access the most recent version at doi:10.1101/gr.6.1.43

LETTER

Structural Organization of the WD Repeat Protein-encoding Gene HIRA in the DiGeorge Syndrome Critical Region of Human Chromosome 22

Stéphanie Lorain,[1] Suzanne Demczuk,[2] Valérie Lamour,[1] Steve Toth,[3] Alain Aurias,[2] Bruce A. Roe,[3] and Marc Lipinski[1,4]

[1]Laboratoire de Biologie des Tumeurs Humaines, Centre National de la Recherche Scientifique URA 1156, Institut Gustave Roussy, 94805 Villejuif CEDEX, France; [2]Laboratoire de Génétique des Tumeurs, Institut National de la Santé et de la Recherche Médicale U 434, Institut Curie, 75231 Paris CEDEX 05, France; [3]Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 78019-0370

The human gene HIRA lies within the smallest critical region for the DiGeorge syndrome (DGS), a haploinsufficiency developmental disorder associated with instertitial deletions in most patients in a juxtacentromeric region of chromosome 22. The HIRA protein sequence can be aligned over its entire length with Hir1 and Hir2, two yeast proteins with a regulatory function in chromatin assembly. The HIRA transcription unit was found to spread over ~100 kb of the DGS critical region. The human transcript is encoded from 25 exons between 59 and 861 bp in size. Domains of highest conservation with Hir1 and Hir2 are encoded from exons 1–11 and 13–25, respectively. The amino- and carboxy-terminal regions of homology are separated from each other by a domain unique to HIRA that is encoded from a single exon. Seven WD repeats are conserved between yeast and man in the amino-terminal region of the HIR proteins. Individual repeats were found to be encoded from one, two, or three exons of the HIRA gene. End sequences have been obtained for all 24 introns, opening the way to PCR amplification of the entire coding sequence starting from genomic DNA. Point mutations can also be sought in 16 of the 24 introns that are readily PCR-amplifiable.

The DiGeorge syndrome (DGS), the Shprintzen, or velo-cardio-facial, syndrome (VCFS), and cases of conotruncal and heart abnormalities represent developmental disorders whose clinical features overlap partially. In accordance, it has been proposed that these disorders be grouped under the CATCH22 (cardiac defect, abnormal facies, thymic hypoplasia, cleft palate, hypocalcemia, chromosome 22q11 deletion) acronym (Wilson et al. 1993). Despite differences in the phenotypes observed in the patients, it appears that tissues affected mostly correspond to derivatives of the third and fourth pharyngeal pouches (Lammer and Opitz 1986). Cytogenetic analysis and molecular genetic studies performed on sporadic patients and, in rare instances, in individuals from families with a history of DGS or related disorders, have indicated that a reduction to hemizygosity for a juxtacentromeric region on the long arm of chromosome 22 could be responsible for these various conditions (Scambler et al. 1991; Carey et al. 1992; Driscoll et al. 1992a,b; Burn et al. 1993; Desmaze et al. 1993; Morrow et al. 1995). DGS, VCFS, and related abnormalities thus contribute examples of haploinsufficiency diseases, a recently recognized group of congenital abnormalities that also includes entities such as the Greig syndrome, the Langer–Giedion syndrome, Hirschprung's disease, the Smith–Magenis syndrome, and the Miller–Dieker syndrome (Fisher and Scambler 1994).

The region of overlap between the chromosomal fragments deleted in different patients defines a critical region for DGS whose centromeric and telomeric boundaries are currently provided by the chromosome 22 breakpoints carried in the X/22-33-11TG somatic cell hybrid, and in the

[4]Corresponding author.
E-MAIL lipinski@citi2.fr; FAX 33 1 45 59 64 94.

GM5878 cell line, respectively (Desmaze et al. 1993). Because the size of the smallest deleted region appears to be consistently several hundred kilobases long, CATCH22 disorders could in fact point to a contiguous gene syndrome, with several genes located within the critical region being collectively responsible for the variability in the phenotypes observed in these patients. Major efforts have been devoted to the identification of possible DGS/VCFS gene candidates. An interesting chromosomal rearrangement, a balanced t(2;22)(q14;q11) translocation, providing a convenient landmark within the region, was originally detected in a patient (ADU) with a partial DGS phenotype (Augusseau et al. 1986). This translocation had been inherited from the proband's mother who was found to exhibit milder symptoms also compatible with a DGS/VCFS diagnosis (Demczuk et al. 1995). The chromosome 22 breakpoint in the translocation propagated in this unique family and in the ADU cell line lies within the DGS smallest critical region. In the last several months, three different groups have identified transcripts encoded from the immediate vicinity of the chromosome 22 breakpoint. A ubiquitous 4.4-kb messenger RNA, encoding a protein with putative adhesive properties, is transcribed from a region immediately telomeric to the ADU breakpoint, but the gene, variously referred to as DGCR2 (Demczuk et al. 1995), IDD (Wadey et al. 1995), or LAN (Budarf et al. 1995), does not appear to be rearranged by the translocation. Using a combination of procedures including cDNA selection, exon trapping, and GRAIL exon prediction from genomic sequences, one of these groups has identified putative coding sequences on each DNA strand in the immediate vicinity of the translocation site in the ADU cell line (Budarf et al. 1995). One of these proposed exons would encompass the chromosomal breakpoint, but the corresponding transcript has not been cloned (Budarf et al. 1995).

A third gene, also in the smallest DGS critical region, lies farther telomeric to the ADU breakpoint (Halford et al. 1993; Lamour et al. 1995). The name Tuple1 had been proposed for the gene because of moderate similarity between the amino acid sequence predicted from a partial cDNA clone and the products of the yeast *TUP1* gene and the fruit fly *Enhancer of split locus* (Halford et al. 1993). The protein alignments we have performed using the amino acid sequence deduced from a complete cDNA have led us to propose that the gene be named HIRA (Lamour et al. 1995), because it encodes a likely homolog of Hir1 and Hir2, two polypeptides with nuclear localization signals but no DNA-binding site that appear to function within a multiprotein complex negatively regulating the transcription of core histone genes in the yeast *Saccharomyces cerevisiae* (Sherwood et al. 1993). In contrast with its carboxy-terminal region, which highly resembles that in Hir2 and similarly lacks any identifiable protein motif, the amino-terminal region of HIRA consists of seven WD repeats, also present in Hir1, and suggests possible interactions with other regulatory proteins (Neer et al. 1994). We have now elucidated the genomic structure of the human gene, which was found to consist of 25 exons spread on a chromosomal segment ~100 kb in length. A relationship is demonstrated between the exon arrangement in the gene and the two distinct regions of similarity to the yeast homologs. In contrast, the seven WD repeats were found to be encoded from one, two, or three exons.

RESULTS AND DISCUSSION

HIRA Complete cDNA and Genomic Organization

The HIRA cDNA sequence originally reported (EMBL accession no. X81844) contained 61 nucleotides upstream of the translation initiation codon (Lamour et al. 1995). The sequence has now been extended further 5′ (accession no. X89887) by incorporation of 159 additional nucleotides present farther upstream in the truncated C5 clone (Halford et al. 1993). RNase protection experiments (data not shown) indicate that the 5′ end of the cDNA sequence corresponds to the main transcription initiation site. The full-length HIRA transcript therefore consists of 4018 nucleotides and a poly(A) tail.

Part of the HIRA gene was known to be contained within the DAC30 cosmid (Lamour et al. 1993). Examination and hybridization of *Eco*RI restriction fragments, along with FISH data colocalizing cosmids DAC30 and 48F8 within the DGS smallest critical region, demonstrated the inclusion of DAC30 into a previously reported cosmid contig (Desmaze et al. 1993), which is represented schematically in Figure 1A. Restriction fragments were hybridized using portions of the HIRA cDNA as probes. The 5′ and 3′ ends of the HIRA cDNA were found to map to the 11-kb

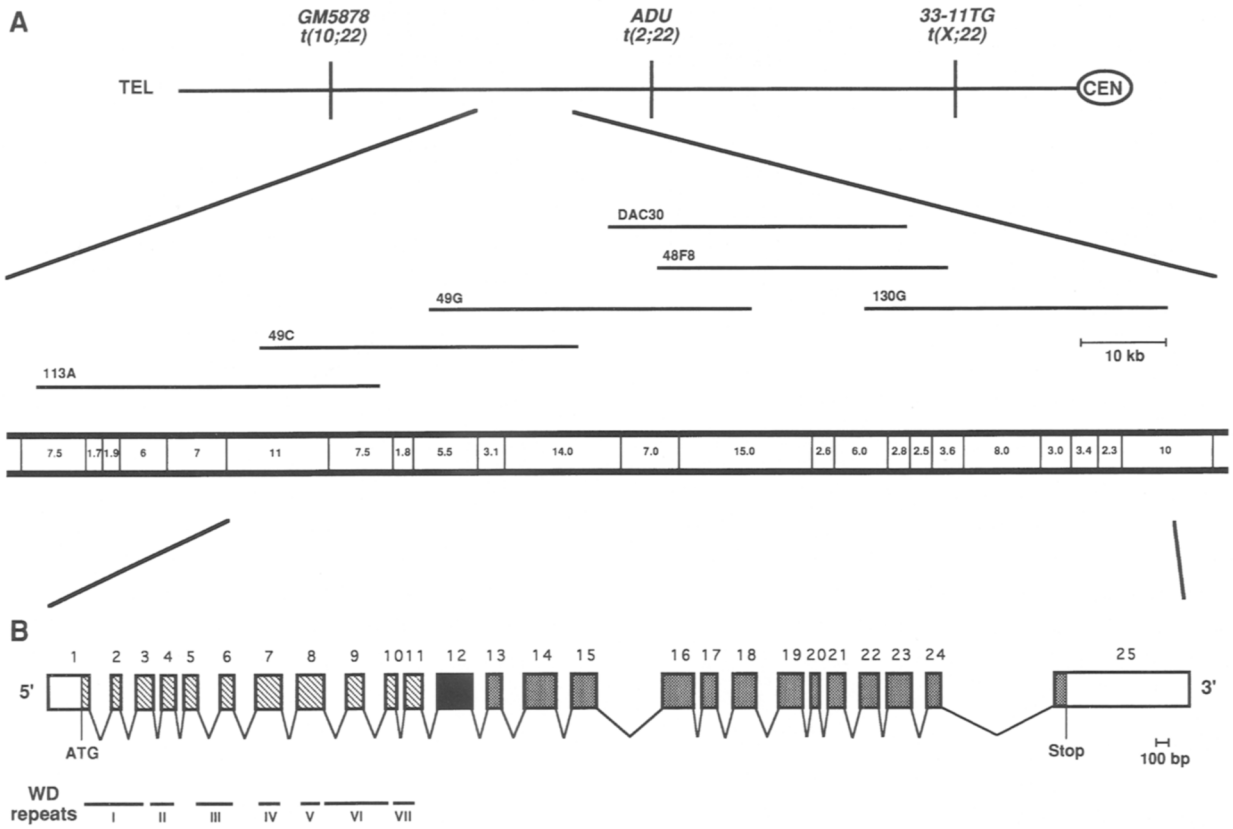

Figure 1 Genomic organization of the HIRA gene in the juxtacentromeric region of chromosome 22q. (*A*) A cosmid contig assembled starting from cosmids DAC30 and 48F8 maps to the region of chromosome 22 that is flanked by the breakpoints in the GM5878 and ADU cell lines. The corresponding *Eco*RI restriction map is shown below. Numbers correspond to the sizes (in kb) of the restriction fragments. (*B*) The gene is transcribed from 25 exons. The translation initiation and stop codons are indicated. HIRA exons (or exon fragments) encoding protein domains of highest similarity to the two yeast homologs Hir1 and Hir2 are represented by hatched and stippled boxes, respectively. Exon 12 (solid box) encodes a protein domain unique to HIRA. The position of the coding sequence for the seven amino-terminal WD repeats is shown below the exon structure.

*Eco*RI fragment in the telomeric 113A cosmid and to the 10-kb *Eco*RI fragment in the centromeric 130G cosmid, respectively (Fig. 1A). From this restriction map, the entire HIRA transcription unit is thus found to stretch over ~100 kb of genomic DNA (Fig. 1A). The exact distance between the DAC30/48F8 contig and that spanning the ADU translocation breakpoint and containing the DGCR2 and DGCR3 genes (Budarf et al. 1995; Demczuk et al. 1995; Wadey et al. 1995) remains to be determined precisely: all attempts to close the gap between the two contigs have been unsuccessful using phage or cosmid libraries, a problem that might be circumvented with the use of new libraries prepared in P1 or BAC vectors.

Restriction fragments containing portions of the HIRA cDNA were identified and cloned. The inserts were sequenced, verifying that the genomic and cDNA sequences were colinear. The complete HIRA cDNA is encoded from 25 exons (Fig. 1B) ranging in size from 59 (exon 20) to 861 (exon 25) bp (Table 1). The translation initiation codon lies within exon 1. The 48 exon–intron boundaries were found to strictly fit the consensus for eukaryotic donor and acceptor splice sites. The length of 16 of the 24 introns, 0.4–2.5 kb in size, was determined by PCR amplification experiments (data not shown) using forward and reverse primers (accession nos. PR00004-PR00019) chosen in the flanking exon sequences.

Relationship Between HIRA Exon Structure and Similarity with the Yeast *HIR1* and *HIR2* Genes

The 5′ one-third (nucleotides 1–1333) of the HIRA cDNA has coding potential for a polypep-

Table 1. Exon Borders and CpG Content—Intron Sizes

Exons								Introns		
no.	CpG (O/E)	size (bp)	Start	5′ sequence		3′ sequence	end	no.	size (kbp)	source
1	1.13	257	1	gatgcggct	...	CAC AAT G	257	1	> 9.3	RM
2	0.79	63	258	GC AAG CCG	...	GGA CAA G	320	2	2.1	PCR
3	0.17	111	321	GG CAG GAT	...	CAC TTA G	431	3	1.2	PCR
4	0.46	91	432	CA TGT GTG	...	CGG GCT AC	522	4	1.4	PCR
5	0.74	95	523	G TAC ATG	...	TCA GGC G	617	5	2.0 - 8.0	RM
6	0.47	96	618	AT GTG ATG	...	TTC CCA G	713	6	2.0 - 8.0	RM
7	0.20	161	714	AA ATT CTA	...	TTT GAT GAG	874	7	2.3	PCR
8	0.44	168	875	TGT GGA GGA	...	ACT GTC GTG	1042	8	4.0 - 9.3	RM
9	0.56	114	1043	AAA TTC AAC	...	TCT GTC TGG	1156	9	< 6.2	RM
10	0.54	71	1157	CTC ACA TGT	...	ATT TCC TG	1227	10	0.7	PCR
11	0.47	106	1228	G ACT CTG	...	GAG GAG AAG	1333	11	2.1	PCR
12	0.43	216	1334	AGC CGC ATT	...	ATC AGG AAG	1549	12	1.9	PCR
13	0.56	86	1550	AAT CTT TTG	...	GAC ACT GG	1635	13	< 8.7	RM
14	0.44	198	1636	G GAC TTC	...	AAA GAC AG	1833	14	2.2	PCR
15	0.61	162	1834	T ATG AAT	...	TGT GGA AAG	1995	15	> 8.6	RM
16	0.42	205	1996	G TTA AAA	...	TCT GTC CAG	2200	16	0.4	PCR
17	0.13	105	2201	TCT CCA GCT	...	ACC CTC CAG	2305	17	1.9	PCR
18	0.55	149	2306	GTC AGC TCC	...	GGC AGC TG	2454	18	2.5	PCR
19	0.64	162	2455	T GAC GTG	...	TCT GTC TG	2616	19	0.6	PCR
20	0.00	59	2617	G GAT GTT	...	CTG GCA G	2675	20	0.4	PCR
21	0.83	106	2676	GA AGT GAT	...	TCC ACA TG	2781	21	1.9	PCR
22	0.30	123	2782	G AAC CTG	...	ACC TCC AA	2904	22	0.6	PCR
23	0.64	164	2905	C TCG GGA	...	AAC GAA G	3068	23	2.1	PCR
24	0.57	89	3069	GG TTT GAA	...	ACA GTA GTG	3157	24	> 17	RM
25	0.39	861	3158	GGT CTG CGG	...	ttcaattac	4018			

Lowercase letters represent untranslated nucleotides. (O/E) The ratio of the number of CpG dinucleotides observed to that expected from the number of cytosines and guanines. Intron sizes were deduced from PCR experiments (PCR) or estimated from the restriction map (RM).

tide of 369 amino acids that is highly similar to a corresponding region in the yeast Hir1 protein: Both proteins contain seven WD repeats, each homologous to its positional counterpart in the other species (Lamour et al. 1995). This region of HIRA was found to be encoded from exons 1 to 11. The yeast *HIR1* gene has been identified as a negative regulator of core histone gene transcription using a genetic screen that also identified *HIR2* (Sherwood et al. 1993). Together, the carboxy-terminal 13 exons of HIRA (exons 13–25) encode a region of significant similarity to the Hir2 protein. In HIRA, these two distinct regions of similarity are separated by an intermediate fragment (amino acids 370–432, encoded by nucleotides 1334–1516) found in neither Hir1 nor Hir2. This transitional region has now been found to be entirely encoded from a single exon (exon 12). Despite the limited similarity between the two yeast transcription factors, Hir1, Hir2, and HIRA can be aligned over their entire lengths (data not shown), suggesting that the three genes might have derived from a common ancestor, with a possible duplication event having occurred in the budding yeast after separation between lower and higher eukaryotic lineages. We

propose that exon 12, which encodes a glutamine-rich peptide, has been inserted in the human transcribed sequence later on during evolution or might have been present in the ancestral transcript, prior to the proposed duplication event, but subsequently lost in the budding yeast. Clues to this question will require the identification of other members in this gene family.

Relationship Between Exon Structure and WD Repeats

In both Hir1 and HIRA, the amino-terminal region of the protein consists of seven WD repeats (Lamour et al. 1995) that are postulated to play a major role in interactions between heterologous proteins (Neer et al. 1994). The cDNA segment encoding the entire WD repeat domain in the human protein terminates precisely at the 3′ end of exon 11 (Fig. 1B). Each WD repeat consists of a core region of 30–45 amino acids, often flanked by conserved dipeptides: a glycine–histidine and a tryptophane–aspartate at the amino- and carboxy-terminal ends, respectively (Neer et al. 1994). In HIRA, no relationship was observed between the exon structure of the gene and the boundaries of the WD repeats: Only two of the seven repeats (repeats IV and V) are entirely encoded within single exons; repeats II, III and VII are encoded from two exons, and repeats I and VI from three (Fig. 1B). Thus, it appears that an important structural aspect has been conserved in yeast and human members of a protein family despite striking differences in the organization of the corresponding genomes.

Nucleotide Composition of HIRA Exons and of the Gene 5′ Region

The overall content in guanines and cytosines, which is 55.8% in the entire HIRA transcript, reaches 80% in exon 1, which is also the only exon with an observed over/expected (O/E) CpG ratio >1 (1.13), whereas it is <0.7 in all other exons except exon 2 (0.79), exon 5 (0.75), and exon 21 (0.83) (Table 1). It is interesting to note that the first partial HIRA cDNA had been originally identified in a library screened with a genomic fragment selected because it appeared to be particularly rich in rare endonuclease sites and thus was suspected to map to the 5′ region of a new gene (Lamour et al. 1993). To our surprise, no cluster of rare endonuclease sites was recognized in this 1.7-kb *Hin*dIII fragment (data not shown), consistent with the observation that it does not map to the 5′, but to a central region, of the HIRA gene which, accordingly, belongs to the category of widely transcribed genes with a CpG island at the start location (Larsen et al. 1992).

A *Pst*I fragment of the 113A cosmid was subcloned and sequenced (EMBL accession no. X91501). It consists of 207 bp of proximal HIRA promoter, followed by exon 1 and by the 5′ end of intron 1. The 5′-most 600 nucleotides (Fig. 2) in this genomic segment are extremely GC-rich (78.3%) with an unusual cluster of 13 rare restriction sites including one *Nar*I (GGCGCC), four *Bss*HII (GCGCGC), three *Xma*III (CGGCCG), three *Sma*I (CCCGGG), and two *Nae*I (GCCGGC) sites. Within this particular stretch of DNA, the observed number of CpG dinucleotides (O) was exactly that expected (E) from the frequency of cytosine and guanine residues, with an O/E ratio of 1.01, indicating that the high number of CpGs in this region have been protected from counterselection during evolution. Finally, and consistent with the ubiquitous transcription of the HIRA gene (Lamour et al. 1995), 207 bp of the proximal HIRA promoter, including three possible binding sites for the Sp1 transcription factor (Kadonaga et al. 1986), was found to lack any TATA or CAAT box (Fig. 2).

```
ctgcaggccccgctcgctgtccgccgtccaggcgancgcgcccctttctgggcgccccgcgc -148
                                                        NarI
ccgggcccgacccgaagcccctgccagcggaggagccgggcgcccagtccgcgtgggg - 88
SmaI
agcggtttcccgccggggcgatttggcaggtgcgcgccgtgacttccggcgttgcccggg - 28
                                BssHII                  SmaI
agccncctgaggaggagcggcgcaggg GATGCGGCTGTGGTGGCGGCGGCGGCGGCGGA   32
                                                        XmaIII
GCGCGGGTGGCGGCTGTGNCGGCGGAGGGGGGCGCGGGCCGGCGATGGCNCGGCGGCCCT   92
                                        NaeI
GAGGGCGCGGGGCGGGCGGCGGCCGCGAGGGCGGGCGGCGCGGGAGGAAGCGGCGGCGGTG  152
                    XmaIII
GCTCCATGGCCCGGGCGCGCTGAGGGACCCGGCGCTCGCCTCAGCCCGGCGGCGGCGGCG  212
          SmaI BssHII                                   Xm
GCCGAACA ATG AAG CTC CTG AAG CCG ACC TGG GTC AAC CAC AAT G     257
aIII      M   K   L   L   K   P   T   W   V   N   H   N

gtgagtgcgcgcagggggtcgcgggaggccgagcccggagtcgggtcggggtctggagtcg  317
      BssHII
agtcgggggcgcgcgcaagtcctgcctgtctcgcccctgacgcccggtacccggtgcccaa  377
      BssHII
gccaagaagccgagcgg                                              393
      NaeI
```

Figure 2 GC-rich segment in the 5′ region of the HIRA gene. Position + 1 corresponds to the 5′ end of the complete cDNA sequence (accession no. X89887). The nucleotide sequence for exon 1 is shown in uppercase letters, with the deduced protein sequence below the coding region. Rare endonuclease sites containing CpG dinucleotides are indicated below the corresponding sequence (underlined). Putative Sp1 binding sites are double-underlined.

HIRA as a DGS Gene Candidate

HIRA is one of several genes identified within the DGS critical region as defined above. Recent findings appear to reinforce the status of HIRA as a prime gene candidate for the syndrome: first, in situ hybridization experiments performed on mouse and chick embryo sections show that the murine and avian homologs of HIRA are expressed at high levels in localized sites, predominantly in rostral regions of the embryos, compatible with implication of the gene product in the early development of the tissues affected in DGS (R. Wadey, C. Roberts, S. Daw, M. Cooke, J. Whiting, E. Lindsay, A. Baldini, I. Dunham, J. Collins, J. Goodship, D. Wilson, J. Burn, H. O'Donell, K. Taylor, S. Halford and P.J. Scambler, pers. comm.); second, molecular cytogenetic studies performed in a typical case of DGS with a chromosome 22 microdeletion have revealed that the proximal boundary of the deleted fragment is distal to the ADU breakpoint (Levy et al. 1995). The latter observation further reduces the shortest region of overlap between DGS deletions to a chromosomal segment that still contains HIRA but no other identified gene.

Finally, a minority of patients with DGS or related disorders do not exhibit a reduction to hemizygosity for markers in the 22q juxtacentromeric region. Accordingly, it has been postulated that other nonchromosome 22 loci could be responsible for DGS-like disorders. It is also possible that the chromosome 22 genes whose reduction to single copy is responsible for most cases of DGS will be found rearranged in a more subtle fashion in patients with two apparently normal chromosomes 22. Whether the HIRA gene plays a role in the etiology of DGS and/or related disorders can now be investigated by a search for single point mutations.

METHODS

Assembly and Characterization of the 48F8 Cosmid Contig

The 48F8 cosmid had been isolated using probe KI-429 (D22S138) (Carey et al. 1990) and mapped within the DGS critical region, between the ADU t(2;22) balanced translocation breakpoint and the t(10;22) translocation breakpoint of the GM5878 cell line (Desmaze et al. 1993). A cosmid contig was constructed starting from 48F8. Hybridizations were performed on two cosmid libraries, LL22NC01 and LL22NC03, constructed in the Lawrence Livermore National Laboratory, together representing eight genome equivalents, using as probes the most extreme *Eco*RI fragments of 48F8 and of subsequently isolated cosmids. The centromere–telomere orientation of the contig was determined by two-color fluorescent in situ hybridization (FISH) on G_0 interphase nuclei and by hybridization of single-copy probes from the contig and neighboring regions on pulsed-field gel electrophoresis blots. The contig covering 170 kb of genomic DNA could not be extended >30 kb toward the centromere, supporting the previous observation that this region is underrepresented in classical genomic libraries (Halford et al. 1993).

Isolation of Genomic Clones and Determination of Exon–intron Borders

Oligonucleotides from the HIRA cDNA sequence and cDNA fragments were used as probes to hybridize endonuclease-digested cosmid DNA. Positive fragments were identified and cloned in the pBluescript II SK + vector (Stratagene, San Diego, CA). Genomic nucleotide sequences were determined by automated sequencing on an ABI 373 instrument (Perkin-Elmer) using flanking vector sequences or cDNA internal oligonucleotides as primers. Exon–intron borders were characterized by alignment between the cDNA and genomic sequences. A 1212-bp *Pst*I fragment from the 113A cosmid was identified by hybridization with a 17-base oligonucleotide representing the 5′ end of the HIRA transcript and entirely sequenced (EMBL accession no. X91501) as described (Bodenteich et al. 1994). Sequencing ambiguities in the GC-rich regions were resolved by including 10% DMSO in the dye-terminator-labeled *Taq* DNA polymerase-catalyzed sequencing reactions.

Intron PCR Amplification

To determine intron sizes, primer oligonucleotides (accession nos. PR00004-PR00019) were chosen from end sequences of the two exons flanking the intron to be amplified by PCR. PCR was performed for 35 cycles, each consisting of 45 sec denaturation at 94°C, 45 sec annealing at 55°C (in the case of introns 2–4, 7, 12, 14, 16, 19–23), or 57°C (introns 10, 11, 17, 18), and 120 sec extension at 72°C, starting with 10 ng of template cosmid DNA prepared by chromatography on Nucleobond AX resin (Macherey-Nagel, France).

ACKNOWLEDGMENTS

S.L. and V.L. were recipients of a fellowship from the Ministère de la Recherche. S.D. is a recipient of fellowships from the Quebec Fonds pour la Formation Chercheurs et l'Aide à la Recherche (FCAR). This work was supported by grants from the Association Française contre les Myopathies (AFM) and the Association pour la Recherche sur le Cancer (ARC) to M.L., from the Groupement de Recherches et d'Etudes sur les Génomes (GREG) to M.L and A.A., and by a National Institutes of Health grant (2 R01 HG00313) to B.A.R. The expert technical assistance of Christine Scamps and Yann Lécluse is warmly acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must there-

fore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Augusseau, S., S. Jouk, P. Jalbert, and M. Prieur. 1986. DiGeorge syndrome and 22q11 rearrangements. *Hum. Genet.* **74:** 206.

Bodenteich, A., S.L. Chissoe, Y.F. Wang, and B.A. Roe. 1994. Shotgun cloning as the strategy of choice to generate templates for high throughput dideoxynucleotide sequencing. In *Automated DNA sequencing and analysis techniques* (ed. M.D. Adams, C. Fields, and C. Venter), pp. 42–50. Academic Press, London, UK.

Budarf, M.L., J. Collins, W. Gong, B. Roe, Z. Wang, L.C. Bailey, B. Sellinger, D. Michaud, D.A. Driscoll, and B.S. Emanuel. 1995. Cloning a balanced translocation associated with DiGeorge syndrome and identification of a disrupted candidate gene. *Nature Genet.* **10:** 269–278.

Burn, J., A. Takao, D. Wilson, I. Cross, K. Momma, R. Wadey, P. Scambler, and J. Goodship. 1993. Conotruncal anomaly face syndrome is associated with a deletion within chromosome 22q11. *J. Med. Genet.* **30:** 822–824.

Carey, A.H., S. Roach, R. Williamson, J.P. Dumanski, M. Nordenskjöld, V.P. Collins, G.A. Rouleau, N. Blin, P. Jalbert, and P.J. Scambler. 1990. Localization of 27 DNA markers to the region of human chromosome 22q11-pter deleted in patients with the DiGeorge syndrome and duplicated in the der22 syndrome. *Genomics* **7:** 299–306.

Carey, A.H., D. Kelly, S. Halford, R. Wadey, D. Wilson, J. Goodship, J. Burn, T. Paul, A.M. Sharkey, J.P. Dumanski, M. Nordenskjöld, R. Williamson, and P.J. Scambler. 1992. Molecular genetic study of the frequency of monosomy 22q11 in Digeorge syndrome. *Am. J. Hum. Genet.* **51:** 964–970.

Demczuk, S., R. Aledo, J. Zucman, O. Delattre, C. Desmaze, L. Dauphinot, P. Jalbert, G.A. Rouleau, G. Thomas, and A. Aurias. 1995. Cloning of a balanced translocation breakpoint in the Digeorge syndrome critical region and isolation of a novel potential adhesion receptor gene in its vicinity. *Hum. Mol. Genet.* **4:** 551–558.

Desmaze, C., M. Prieur, F. Amblard, M. Aïkem, F. LeDeist, S. Demczuk, J. Zucman et al. 1993. Physical mapping by FISH of the DiGeorge critical region (DGCR): Involvement of the region in familial cases. *Am. J. Hum. Genet.* **53:** 1239–1249.

Driscoll, D.A., M. Budarf, and B.S. Emanuel. 1992a. A genetic etiology for DiGeorge syndrome: Consistent deletions and microdeletions of 22q11. *Am. J. Hum. Genet.* **50:** 924–933.

Driscoll, D.A., N.B. Spinner, M.L. Budarf, D.M. McDonald-McGinn, E.H. Zackai, R.B. Goldberg, R.J. Shprintzen, H.M. Saal, J. Zonana, M.C. Jones, J.T. Mascarello, and B.S. Emanuel. 1992b. Deletions and microdeletions of 22q11.2 in Velo-Cardio-Facial syndrome. *Am. J. Med. Genet.* **44:** 261–268.

Fisher, E. and P. Scambler. 1994. Human haploinsufficiency—One for sorrow, two for joy. *Nature Genet.* **7:** 5–7.

Halford, S., R. Wadey, C. Roberts, S.C.M. Daw, J.A. Whiting, H. O'Donnell, I. Dunham et al. 1993. Isolation of a putative transcriptional regulator from the region of 22q11 deleted in DiGeorge syndrome, Shprintzen syndrome and familial congenital heart disease. *Hum. Mol. Genet.* **2:** 2099–2107.

Kadonaga, J.T., K.A. Jones, and R. Tjian. 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. *Trends Biochem. Sci.* **11:** 20–23.

Lammer, E.J. and J.M. Opitz. 1986. The DiGeorge anomaly as a developmental field defect. *Am. J. Med. Genet.* (Suppl.) **2:** 113–127.

Lamour, V., N. Lévy, C. Desmaze, V. Baud, Y. Lécluse, O. Delattre, A. Bernheim, G. Thomas, A. Aurias, and M. Lipinski. 1993. Isolation of cosmids and fetal brain cDNAs from the proximal long arm or human chromosome 22. *Hum. Mol. Genet.* **2:** 535–540.

Lamour, V., Y. Lécluse, C. Desmaze, M. Spector, M. Bodescot, A. Aurias, M.A. Osley, and M. Lipinski. 1995. A human homolog of the *S. cerevisiae* HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. *Hum. Mol. Genet.* **4:** 791–799.

Larsen, F., G. Gundersen, R. Lopez, and H. Prydz. 1992. CpG islands as gene markers in the human genome. *Genomics* **13:** 1095–1107.

Levy, A., S. Demczuk, A. Aurias, D. Depetris, M.-G. Mattei, and N. Philip. 1995. Insterstitial microdeletion excluding the ADU breakpoint in a patient with DiGeorge syndrome. *Hum. Mol. Genet.* **4:** 2417–2419.

Morrow, B., R. Goldberg, C. Carlson, R. Das Gupta, H. Sirotkin, J. Collins, I. Dunham, H. O'Donnell, P. Scambler, R. Shprintzen, and R. Kucherlapati. 1995. Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am. J. Hum. Genet.* **56:** 1391–1403.

Neer, E.J., C.J. Schmidt, R. Nambudripad, and T.F. Smith. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* **371:** 297–300.

Scambler, P.J., A.H. Carey, R.K.H. Wyse, S. Roach, J.P. Dumanski, M. Nordenskjöld, and R. Williamson. 1991. Microdeletions within 22q11 associated with sporadic and familial DiGeorge Syndrome. *Genomics* **10:** 201–206.

Sherwood, P.W., S.V. Tsang, and M.A. Osley. 1993. Characterization of HIR1 and HIR2, two genes required

for regulation of histone gene transcription in Saccharomyces cerevisiae. *Mol. Cell Biol.* **13:** 28–38.

Wadey, R., S. Daw, C. Taylor, U. Atif, S. Kamath, S. Halford, H. O'Donnell, D. Wilson, J. Goodship, J. Burn, and P. Scambler. 1995. Isolation of a gene encoding an integral membrane protein from the vicinity of a balanced translocation breakpoint associated with DiGeorge syndrome. *Hum. Mol. Genet.* **4:** 1027–1033.

Wilson, D., J. Burn, P. Scambler, and J. Goodship. 1993. DiGeorge syndrome: part of CATCH 22. *J. Med. Genet.* **30:** 852–856.

Received September 25, 1995; accepted in revised form December 21, 1995.